

Geometry of Thermodynamic States

Dorje C. Brody* and Lane P. Hughston†

**Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, Silver Street, Cambridge CB3 9EW U.K.*

*† Merrill Lynch International, 25 Ropemaker Street, London EC2Y 9LY U.K.
and King's College London, The Strand, London WC2R 2LS, U.K.*

(February 9, 2008)

A novel geometric formalism for statistical estimation is applied here to the canonical distribution of classical statistical mechanics. In this scheme thermodynamic states, or equivalently, statistical mechanical states, can be characterised concisely in terms of the geometry of a submanifold \mathcal{M} of the unit sphere \mathcal{S} in a real Hilbert space \mathcal{H} . The measurement of a thermodynamic variable then corresponds to the reduction of a state vector in \mathcal{H} to an eigenstate, where the transition probability is the Boltzmann weight. We derive a set of uncertainty relations for conjugate thermodynamic variables in the equilibrium thermodynamic states. These follow as a consequence of a striking thermodynamic analogue of the Anandan-Aharonov relations in quantum mechanics. As a result we are able to provide a resolution to the controversy surrounding the status of ‘temperature fluctuations’ in the canonical ensemble. By consideration of the curvature of the thermodynamic trajectory in its state space we are then able to derive a series of higher order variance bounds, which we calculate explicitly to second order.

PACS Numbers : 05.20.Gg, 05.70.Ce, 02.40.Ky, 02.50.Cw

By a statistical model \mathcal{M} we mean a family of probability distributions characterised by a set of parameters known as the parameter space. This possesses natural geometrical properties induced by the embedding of the family of probability distributions in the space of square-integrable functions on the relevant sample space. More precisely, by consideration of the parameterised square-root density function we can regard the space \mathcal{M} as a submanifold of the unit sphere \mathcal{S} in a real Hilbert space \mathcal{H} . Therefore, \mathcal{H} embodies the state space of the system, and the properties of the statistical model can be described in terms of the embedding of \mathcal{M} in \mathcal{H} . The geometry thus arising possesses a natural Riemannian metric, the Fisher-Rao metric, and as a consequence of this the powerful tools of differential geometry can be applied to various aspects of statistical inference.

To what extent is this methodology applicable to statistical physics? In the present Letter, we focus on the statistical model \mathcal{M} that constitutes the state space of classical statistical mechanics, for which the corresponding probability density is given by the Gibbs distribution. By taking the square root of this density, we can map, for each value of the parameter, the associated probability density to a vector in \mathcal{H} . We are thus led to formulate a classical theory of measurement and statistical estimation in the language of Hilbert space geometry, with applications to classical and quantum statistical mechanics. This formulation of statistical theory ties up with the generalised probability theory due to Segal [1]. Our aim here, however, is to formulate matters in a geometric manner, in such a way that the statistical operations associated with estimation problems become more transparent.

As a consequence, we find that a classical thermodynamic state, in the energy representation, can be expressed as a real superposition of energy eigenstates, each having the square-root of the corresponding Boltzmann weight as coefficient. The temperature estimation in such equilibrium states suffers from an intrinsic uncertainty, leading to what we might call a thermodynamic uncertainty relation. The source of the uncertainty can be said to be thermal noise, which is analogous to the quantum noise that gives rise to the Heisenberg uncertainty relations. Indeed, quantum mechanical Schrödinger trajectories can be obtained by Wick rotating the corresponding classical thermal trajectories on \mathcal{S} , and hence we can enquire to what extent standard quantum mechanical relations have analogues in a classical thermodynamic system. We shall demonstrate, for example, that the classical relationship between the heat capacity of a system in thermal equilibrium and the energy variance in the associated Gibbs distribution is in direct correspondence with the Anandan-Aharonov relation [2] in quantum mechanics, which associates the velocity of a quantum mechanical state trajectory with the energy uncertainty along it. From there we are led to a series of higher order variance bounds on the temperature measurement, which also have quantum mechanical analogues [3].

Consider a real Hilbert space \mathcal{H} with a symmetric inner product g_{ab} . A probability density function $p(x)$ can be mapped into \mathcal{H} by taking the square-root $\psi(x) = \sqrt{p(x)}$, which is denoted by a vector ψ^a in \mathcal{H} . The normalisation condition $\int (\psi(x))^2 dx = 1$, written $g_{ab}\psi^a\psi^b = 1$, indicates that ψ^a lies on the unit sphere \mathcal{S} in \mathcal{H} . Since a probability density function is nonnegative, the image of the mapping $f : p(x) \rightarrow \psi(x)$ is the intersection $\mathcal{S}_+ = \mathcal{S} \cap \mathcal{H}_+$ of the sphere \mathcal{S} in \mathcal{H} with the convex cone \mathcal{H}_+ formed by the totality of quadratically integrable nonnegative functions. We call ψ^a the state vector of the corresponding probability density $p(x)$.

A typical random variable is represented on \mathcal{H} by a symmetric tensor X_{ab} , whose expectation in a normalised state ψ^a is $E_\psi[X] = X_{ab}\psi^a\psi^b$. Similarly, the expectation of its square is $X_{ac}X_b^c\psi^a\psi^b$. The variance of X_{ab} in the state ψ^a is therefore $\text{Var}_\psi[X] = \tilde{X}_{ac}\tilde{X}_b^c\psi^a\psi^b$, where $\tilde{X}_{ab} = X_{ab} - g_{ab}E_\psi[X]$ represents the deviation of X_{ab} from its mean in the state ψ^a .

Suppose we have a family of probability distributions that are conditioned to a set of parameters θ , with density function $p(x, \theta)$. Then, for each value of θ we obtain a corresponding point on \mathcal{S} given by $\psi^a(\theta)$. This conditioning is characterised by the specification of a submanifold \mathcal{M} in \mathcal{S} . Assuming $\psi^a(\theta)$ has continuous second derivatives, we find that \mathcal{M} is a Riemannian manifold, with the Fisher-Rao metric $G_{ij} = 4g_{ab}\partial_i\psi^a\partial_j\psi^b$, where $\partial_i = \partial/\partial\theta^i$. This is the metric induced on \mathcal{M} by the spherical geometry of \mathcal{S} .

In a statistical mechanical context, the parametrised family of probability distribution takes the form of the Gibbs measure

$$p(x, \theta) = q(x) \exp \left[- \sum_j \theta^j H_j(x) - W_\theta \right], \quad (1)$$

where the variable x ranges over the configuration space, $H_j(x)$ represents the form of the energy, W_θ is a normalisation factor, and $q(x)$ determines the distribution at $\theta^j = 0$. Our goal is to formulate a Hilbert space characterisation of this distribution. In fact, it can be shown [3] that the state vector $\psi^a(\theta)$ in \mathcal{H} corresponding to the Gibbs distribution (1) satisfies the differential equation

$$\frac{\partial \psi^a}{\partial \theta^j} = -\frac{1}{2} \tilde{H}_{jb}^a \psi^b, \quad (2)$$

where $\tilde{H}_{jab} = H_{jab} - g_{ab} E_\psi[H_j]$. The solution of this equation is

$$\psi^a(\theta) = \exp \left[-\frac{1}{2} \left(\sum_j \theta^j H_{jb}^a + \tilde{W}_\theta \delta_b^a \right) \right] q^b, \quad (3)$$

where $\tilde{W}_\theta = W_\theta - W_0$ and $q^a = \psi^a(0)$ is the prescribed distribution at $\theta^j = 0$. The Fisher-Rao metric for the parameter space \mathcal{M} of the Gibbs distribution is $G_{ij} = \partial_i \partial_j W_\theta$. Thus, for models in statistical mechanics where the normalisation is known, or equivalently, for which the partition function is specified, one can study the geometry of \mathcal{M} directly by use of this expression [4].

We are now in a position to set up a microscopic theory of measurement for thermodynamic systems. For simplicity, we consider a one parameter family of thermal states, letting $\beta = 1/k_B T$ denote the usual inverse-temperature parameter and H_{ab} the symmetric Hamiltonian for the system. For the state vector we have

$$\psi^a(\beta) = \exp \left[-\frac{1}{2} (\beta H_b^a + \tilde{W}_\beta \delta_b^a) \right] q^b, \quad (4)$$

and for each value of the temperature we find a point on \mathcal{M} in \mathcal{S} . To be more specific, we have a unit sphere \mathcal{S} in \mathcal{H} , whose axes label the configurations of the system, each of which has a definite energy. Suppose we let u_k^a denote an orthonormal basis of vectors in \mathcal{H} . Here, the index k labels all the points in the phase space of the given statistical system. In other words, for each point in phase space we have a corresponding basis vector u_k^a in \mathcal{H} for some value of k . With this choice of basis, a classical thermal state $\psi^a(\beta)$ can be expressed as a superposition

$$\psi^a(\beta) = e^{-\frac{1}{2} W_\beta} \sum_k e^{-\frac{1}{2} \beta E_k} u_k^a, \quad (5)$$

where E_k is the energy for k -th configuration, and thus $\exp[W_\beta] = \sum_k \exp(-\beta E_k)$ is the partition function. The index k is formal here in the sense that the summation may if appropriate be replaced by an integration. By comparing equations (4) and (5), we find that the $\beta = 0$ thermal state q^a is given by

$$q^a = e^{-\frac{1}{2} W_0} \sum_k u_k^a, \quad (6)$$

which corresponds to the centre point in \mathcal{S}_+ . This relation reflects the fact that all configurations are equally likely to occur at infinite temperature, with probability $\exp(-W_0)$. The state trajectory $\psi^a(\beta)$ thus commences at the centre point q^a , and follows a curve on \mathcal{S} generated by the Hamiltonian H_{ab} , for which the curvature is

$$K_\psi(\beta) = \frac{\langle \tilde{H}^4 \rangle}{\langle \tilde{H}^2 \rangle^2} - \frac{\langle \tilde{H}^3 \rangle^2}{\langle \tilde{H}^2 \rangle^3} - 1, \quad (7)$$

where $\langle \tilde{H}^n \rangle$ denotes the n -th central moment of the observable H_{ab} . Here as usual the curvature of the curve $\psi^a(\beta)$, which is necessarily positive, is defined by $K_\psi(\beta) =$

$g_{ab}\psi_2^a\psi_2^b/(g_{ab}\psi_1^a\psi_1^b)^2$, where $\psi_2^a = \ddot{\psi}^a - \dot{\psi}^a\ddot{\psi}^b\dot{\psi}_b/\dot{\psi}^c\dot{\psi}_c - \psi^a\ddot{\psi}^b\dot{\psi}_b$ is the ‘acceleration’ vector along $\psi^a(\beta)$, and $\psi_1^a = \dot{\psi}^a$ is the ‘velocity’ (the dot denotes $\partial/\partial\beta$). The acceleration satisfies $\psi_2^a\psi_a = 0$ and $\psi_2^a\psi_{1a} = 0$. Note that $\psi_1^a\psi_a = 0$ since $\psi^a(\beta)$ lies on \mathcal{S} .

If the dimension of \mathcal{H} is infinite, the statistical model may exhibit a phase transition at a critical point β_c . Then the curve proliferates into L distinct curves, where L is the multiplicity of the ground state degeneracy. Thus, for the thermal states it is important to fix the initial condition at $\beta = 0$, since any other point can be ambiguous. Physically, this ambiguity reflects the various coexisting phases allowed at the critical point. In particular, if the transition is of second order, the curvature is singular at β_c . It follows from the expression (7) that the scaling behaviour of K_ψ around β_c is given by $K_\psi \sim |\hat{\beta}|^{-\kappa}$, where $\hat{\beta} = \beta_c/\beta - 1$ is the reduced temperature and $\kappa = 2 - \alpha$ in terms of conventional critical exponents. The standard relation $2 - \alpha = d\nu$ indicates that the curvature scales like correlation volume [4].

The simplest model for a measurement of the state can be described by projecting out a point in phase space. The resulting probability for observing the state u_k^a for some value of k is thus given by the Boltzmann weight

$$p_k = (g_{ab}\psi^a u_k^b)^2 = e^{-\beta E_k - W_\beta} . \quad (8)$$

This model can be extended to incorporate probability operator valued measures. For instance, suppose we are interested in the measurement of an observable X_{ab} for a state $\psi^a(\beta)$. Then, the probability density for the measurement outcome x is the expectation $p(x, \beta) = \Pi_{ab}(X, x)\psi^a\psi^b$ of the projection operator

$$\Pi_b^a(X, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp[i\lambda(X_b^a - x\delta_b^a)] d\lambda .$$

The choice of Π_{ab} can also include nonorthogonal resolution operators.

Now let us turn to the measurement problem for thermal states. Our intention is to study the uncertainties arising in making inferences from the measurement outcomes. In general, we may wish to estimate the value of a function of an unknown parameter, such as internal energy or magnetic susceptibility. We shall consider, in particular, the case when we estimate the temperature. Suppose that B_{ab} is an unbiased estimator for the parameter β , so $B_{ab}\psi^a\psi^b/g_{cd}\psi^c\psi^d = \beta$. Then, the variance in estimating β can be expressed [5,6] by the geometrical relation

$$\text{Var}_\psi[B] = \frac{1}{4}g^{ab}\nabla_a\beta\nabla_b\beta , \quad (9)$$

on the unit sphere \mathcal{S} , where $\nabla_a\beta = \partial\beta/\partial\psi^a$ is the gradient of the temperature estimate β . The essence of this relation can be understood as follows. First, recall that β is the expectation of the operator B_{ab} in the state $\psi^a(\beta)$. Suppose that the state changes rapidly as β changes. Then, the variance in estimating β is small, and indeed, this is given by the squared magnitude of the ‘functional derivative’ of β with respect to the state ψ^a . On the other hand, if the state does not change significantly as β changes, then the measurement outcome of an observable is less conclusive in determining the value of β . A crude example is as follows. Suppose we infer the value of the temperature for a magnetic system from measurements of the magnetisation. If the measurement outcome is, say, close to zero, then the temperature can be any value above the Curie point, and the variance is large.

The squared length of the gradient vector $\nabla_a \beta$ can be expressed as a sum of squares of orthogonal components. To this end, we choose a new set of orthogonal basis vectors given by the state ψ^a and its higher order derivatives. If we let ψ_n^a denote ψ^a for $n = 0$, and for $n > 0$ the component of the derivative $\partial^n \psi^a / \partial \beta^n$ orthogonal to the state ψ^a and its lower order derivatives, then our orthonormal vectors are given by $\hat{\psi}_n^a = \psi_n^a (g_{bc} \psi_n^b \psi_n^c)^{-1/2}$ for $n = 0, 1, 2, \dots$. The reason for choosing this set instead of the original basis u_k^a is for computational simplicity. With this choice of orthonormal vectors, we find that the variance of the estimator B satisfies the inequality

$$\text{Var}_\psi[B] \geq \sum_n \frac{(\tilde{B}_{ab} \psi_n^a \psi_n^b)^2}{g_{cd} \psi_n^c \psi_n^d}, \quad (10)$$

for any range of the index n . This follows as a consequence of the fact that the squared magnitude of the vector $\frac{1}{2} \nabla_a \beta = \tilde{B}_{ab} \psi^b$ is necessarily greater than or equal to the sum of the squares of its projections onto the basis vectors given by $\hat{\psi}_n^a$ for the specified range of n .

In particular, for $n = 1$ we have $B_{ab} \psi_1^a \psi^b = \frac{1}{2}$ on account of the relation $B_{ab} \psi^a \psi^b = \beta$, and $g_{ab} \psi_1^a \psi_1^b = \frac{1}{4} \Delta H^2$, which follows from the differential equation $\partial \psi^a / \partial \beta = -\frac{1}{2} \tilde{H}_b^a \psi^b$. Therefore, as a consequence of equation (10), if we write $\text{Var}_\psi[B] = \Delta \beta^2$, we find for $n = 1$ the following thermodynamic uncertainty relation:

$$\Delta \beta^2 \Delta H^2 \geq 1 \quad (11)$$

which is valid along the trajectory \mathcal{M} consisting of the thermal equilibrium states $\psi^a(\beta)$. Interestingly, Landau and Lifshitz have obtained an inequality of this kind on the basis of temperature fluctuations. Kittel and Kroemer argue, on the other hand, that such an inequality is meaningless, since the temperature is a fixed constant by definition in the canonical distribution. The framework we have presented provides a mathematically and physically consistent solution to this long standing point of controversy [7]. That is, β is indeed a fixed constant for a canonical ensemble, which does not fluctuate. However, for a given equilibrium system, if we wish to find the actual value of β , there is an inevitable uncertainty associated with our estimation, characterised by (11).

While the variance bounds in (10) formally depend on the specific choice of the estimator B_{ab} , we find, remarkably, that in the case of thermal states, these bounds are systematically independent of the estimator B . For example, for $n = 2$, the acceleration vector ψ_2^a is given by the expression

$$\psi_2^a = \frac{1}{4} \left(\tilde{H}_b^a \tilde{H}_c^b \psi^c - \frac{\langle \tilde{H}^3 \rangle}{\langle \tilde{H}^2 \rangle} \tilde{H}_b^a \psi^b - \langle \tilde{H}^2 \rangle \psi^a \right).$$

To value the corresponding correction term in (10) we note that $\tilde{B}_{ab} \psi_2^a \psi^b = B_{ab} \psi_2^a \psi^b$ on account of the orthogonality $g_{ab} \psi_2^a \psi^b = 0$. Then, by differentiating the relation $B_{ab} \psi^a \psi^b = \beta$, we obtain $B_{ab} \tilde{H}_c^b \psi^a \psi^c = -1$. Taking a second derivative we conclude that $B_{ab} \tilde{H}_c^a \tilde{H}_d^b \psi^c \psi^d = \beta \langle \tilde{H}^2 \rangle$, providing that B_{ab} commutes with H_{ab} , which is the case for classical thermodynamic variables. Substitution of these relations into (10) yields a sharper thermodynamic uncertainty relation

$$\Delta \beta^2 \Delta H^2 \geq 1 + \frac{\langle \tilde{H}^3 \rangle^2}{\langle \tilde{H}^2 \rangle^3 K_\psi}, \quad (12)$$

where the expression for the curvature K_ψ is given in (7). It is clear that, by consideration of other equilibrium distributions such as grand canonical or P - T distributions, analogous relations can be derived, e.g., for chemical potential $\mu/k_B T$ and particle number N , or pressure $P/k_B T$ and volume V . The equality in (11) holds if the energy expectation is proportional to the inverse temperature β . It is also interesting to observe that by virtue of the formula $2B_{ab}\psi_1^a\psi_1^b = 1$, these conjugate variables satisfy the covariance relation $E_\psi[BH] - E_\psi[B]E_\psi[H] = -1$, which can be expressed in the form of the anticommutation relation $\{B, \tilde{H}\}_\psi = -1$, valid in expectation along the thermal trajectory.

We note, incidentally, that the Fisher-Rao metric $G = 4g_{ab}\psi_1^a\psi_1^b$ in this case is given by $\Delta H^2 = T^2 C$, where C is the heat capacity. For systems exhibiting second order phase transitions, the heat capacity diverges at the critical point. Therefore, the temperature uncertainty $\Delta\beta^2$ can be made small in the vicinity of such critical points. This is due to the sensitivity of the state near critical points. Conversely, it is clear that temperature estimation becomes difficult for values of β far from the critical value β_c . The relation $G = \Delta H^2$ is the thermodynamic counterpart of the Anandan-Aharonov relation [2] in geometric quantum mechanics. However, unlike the quantum case where G is constant along Schrödinger trajectories, for thermal trajectories G depends upon the parameter β . In particular, we find that $\partial_\beta G = 2\langle \tilde{H}^3 \rangle$, which shows that for the higher order correction given in (12) to be nontrivial, it suffices that the bound in (11) should not be saturated.

In the foregoing sketch of our theory of statistical measurement, we have observed that many standard quantum mechanical operations are already present at an essentially classical level of probabilistic reasoning. This is surprising, since the general view in physics is that the Hilbert space structure associated with the space of states in nature is special to quantum theory, and has no analogue in classical probability theory and statistics. Therefore, we can ask to what extent quantum theory is distinguished from the classical probability theory discussed above. From the viewpoint of Hilbert space geometry, indeed, the quantum theory can be regarded in a certain sense as a special case of the generalised statistical theory. That is, the theory outlined above can be specialised to the space of quantum states by introducing a compatible complex structure J_b^a on the underlying real Hilbert space \mathcal{H} . In this case, the state trajectory $\xi^a(t)$, having the time t as parameter, satisfies the Schrödinger equation $\partial\xi^a/\partial t = J_b^a \tilde{H}_c^b \xi^c$ and substitution of this into equation (10), for $n = 1$, yields the quantum mechanical uncertainty relation $\Delta t^2 \Delta H^2 \geq 1/4$.

The constructions given above for the thermodynamic equilibrium states $\psi^a(\beta)$ and the Schrödinger trajectories $\xi^a(t)$ illustrate two examples of probabilistic theories that can be analysed within the framework of Hilbert space geometry. For nonequilibrium systems, on the other hand, we may need to consider statistical dynamics [8], in which case the differential equation for the state ψ^a has to be modified appropriately. Once a model has been chosen, the formalism can be applied to study the geometrical and statistical properties of these systems.

Our approach has been to view the thermodynamic state space as an example that can be put forth within the framework of a generalised statistics, with an emphasis on the geometric structure of the underlying real Hilbert space. This generalised probability theory might at first appear to be formulated by way of an *ad-hoc* analogy with quantum theory. However, as we proceed, we find that the Hilbert space structure associated with the probability distributions is indeed the fundamental construction. In this formulation, various

concepts in statistical studies, such as the notion of uncertainty, have precise geometric characterisations, hence allowing a transparent understanding of the underlying physics of the given model. We note that the higher order correction in (12) can easily be calculated for specific models. For example, for an N -spin Ising chain, we find that the correction term is given by $2 \sinh^2(\beta J)/(N - 1)$, where J is the exchange integral.

The authors acknowledge their gratitude to B. K. Meister and R. F. Streater for useful discussions. DCB is also grateful to PPARC for financial support.

* Electronic address: d.brody@damtp.cam.ac.uk

† Electronic address: lane@ml.com

1. I.E. Segal, Ann. Math. **48**, 930 (1947).
2. J. Anandan and Y. Aharonov, Phys. Rev. Lett. **65**, 1697 (1990).
3. D.C. Brody and L.P. Hughston, Phys. Rev. Lett. **77**, 2851 (1996).
4. Ruppiner, Rev. Mod. Phys., **67**, 605 (1995); D. Brody and N. Rivier, Phys. Rev. E, **51**, 1006 (1995); H. Janyszek and R. Mrugala, Phys. Rev. A, **39**, 6515 (1989).
5. D.C. Brody and L.P. Hughston, “Statistical Geometry”, Preprint IC/TP/95-96/42, gr-qc/9701051.
6. D.C. Brody and L.P. Hughston, in *Geometric Issues in the Foundations of Science*, edited by S.A. Huggett, et. al. (Oxford University Press, Oxford 1997).
7. L.D. Landau and E.M. Lifshitz, *Statistical Physics*, (Pergamon Press, Oxford 1980); C. Kittel and H. Kroemer, *Thermal Physics*, (W.H. Freeman, San Francisco 1980).
8. R.F. Streater, *Statistical Dynamics*, (Imperial College Press, London 1995).